TS EAS

Monday 3rd of August

6.00 am PDT / 7.00 am MDT / 8.00 am CDT / 9.00 am EDT / 2.00pm BST / 3.00 pm CEST / 11.00pm AEST

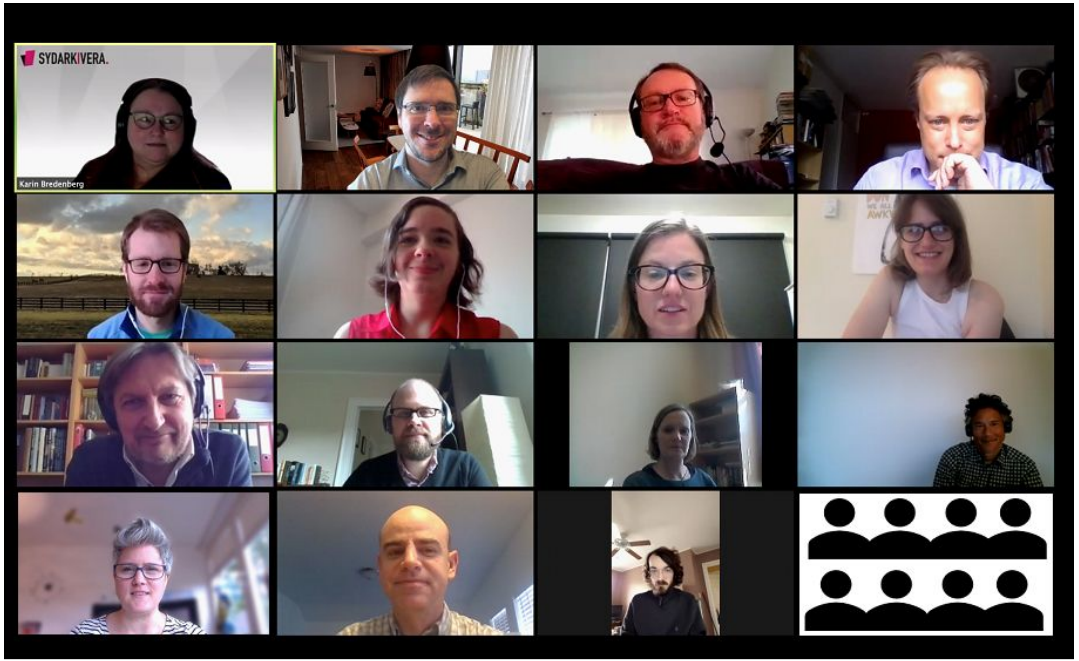
Welcome to this week!

5 days, 5 meetings!

Housekeeping

- Cameras
- A short break when it fits if needed
- No recordings....
- All slides are in one presentation
    - We will figure out how we do the sharing of screen easiest during today, take over screen/tell to turn/switch presenter

2019-2020



Notes

- Jess is lead
- Collaborative!

Slack

- We (Mark) have set up a Slack channel
- For short quick messages
- You need an invite (Mark)
- Can be used for the teams.
    - One channel per team, team lead decision
- https://slack.com/intl/en-se/

Last chance for voting!

Mark’s email from the 2nd of July.

Do we follow a number of Camels in a row, eat a kebab or just ignore the shift key on our keyboard?

Open for all meeting

- We'll have an open for all meeting on the 13th.
  - Need to sign up for getting the connection info
  - https://zoom.us/meeting/register/tJEsd-Cupz8pHdfjGcIf9XNholrPhQ9rogbB
- A chance for the members of SAA to give comments outside of an RFC
- Info sent by Adrian and Cory to relevant lists

Schema design principles

Mark

Background

Two schemas, managed in two different GitHub repositories, with two vastly different publication processes inherited from separate groups:

- EAC-CPF:
  - RNG -> manual conversion to XSD, with an additional manual change to include the XLink schema.
  - Also a major difference between the RNG and XSD variants.
- EAD3:
  - Semi-automated process, with roughly 10 steps to convert an RNG “driver” schema to 6 different deliverables.

Now...

We have a single committee, and a single Schema development team.

Four goals:

- Unite the schema-publication process
- Ensure that there are no validation-based differences between the deliverables
- Automate the publication process
- Add tests to ensure valid example documents remain valid and vice versa

More on the upcoming schema publication process later, but next let's discuss general principles that can help guide future decisions.

Draft principles

TS-EAS Schema Design Principles (July version)

File: TS-EAS Schema Design Principles (July version).pdf

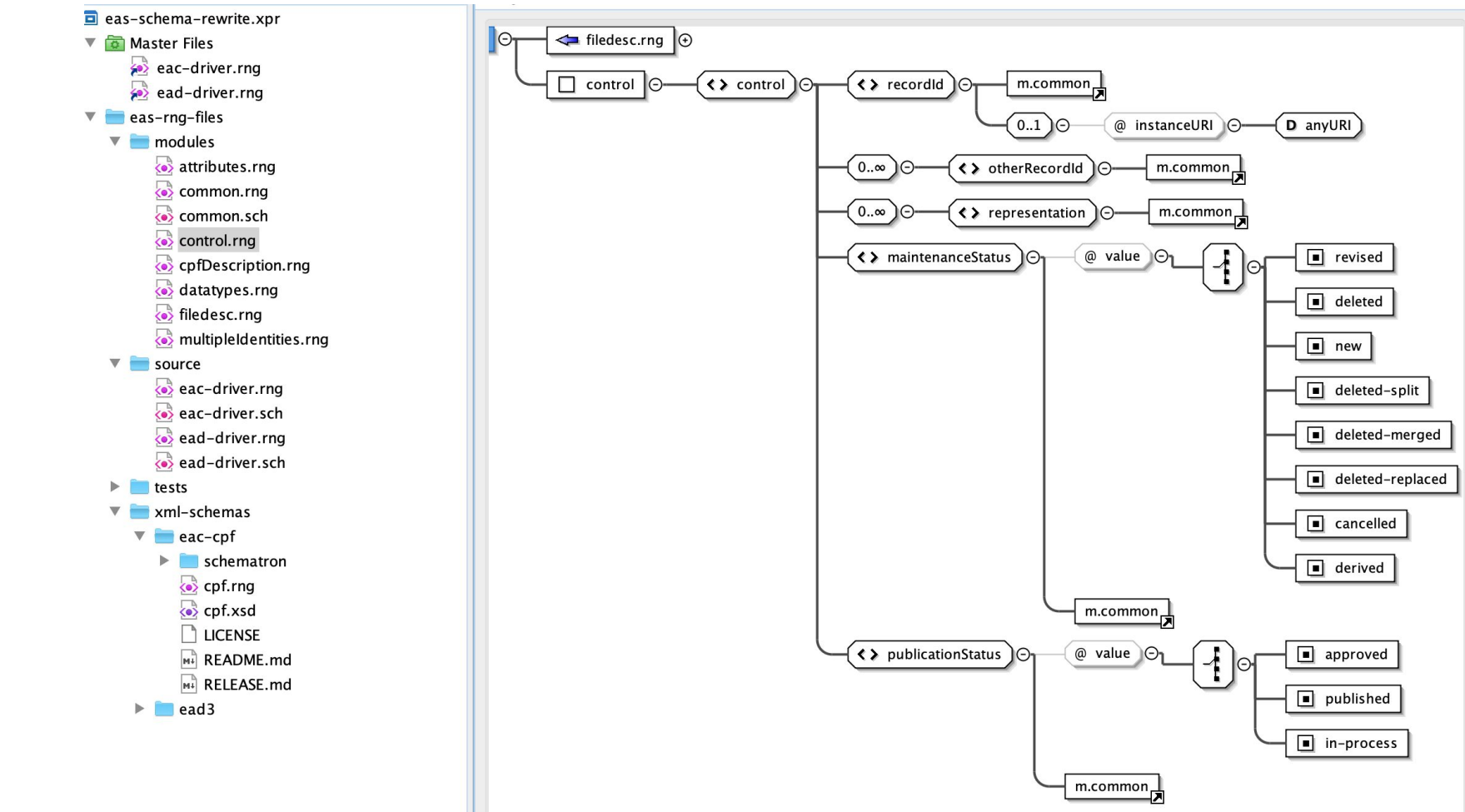
Schema publication changes

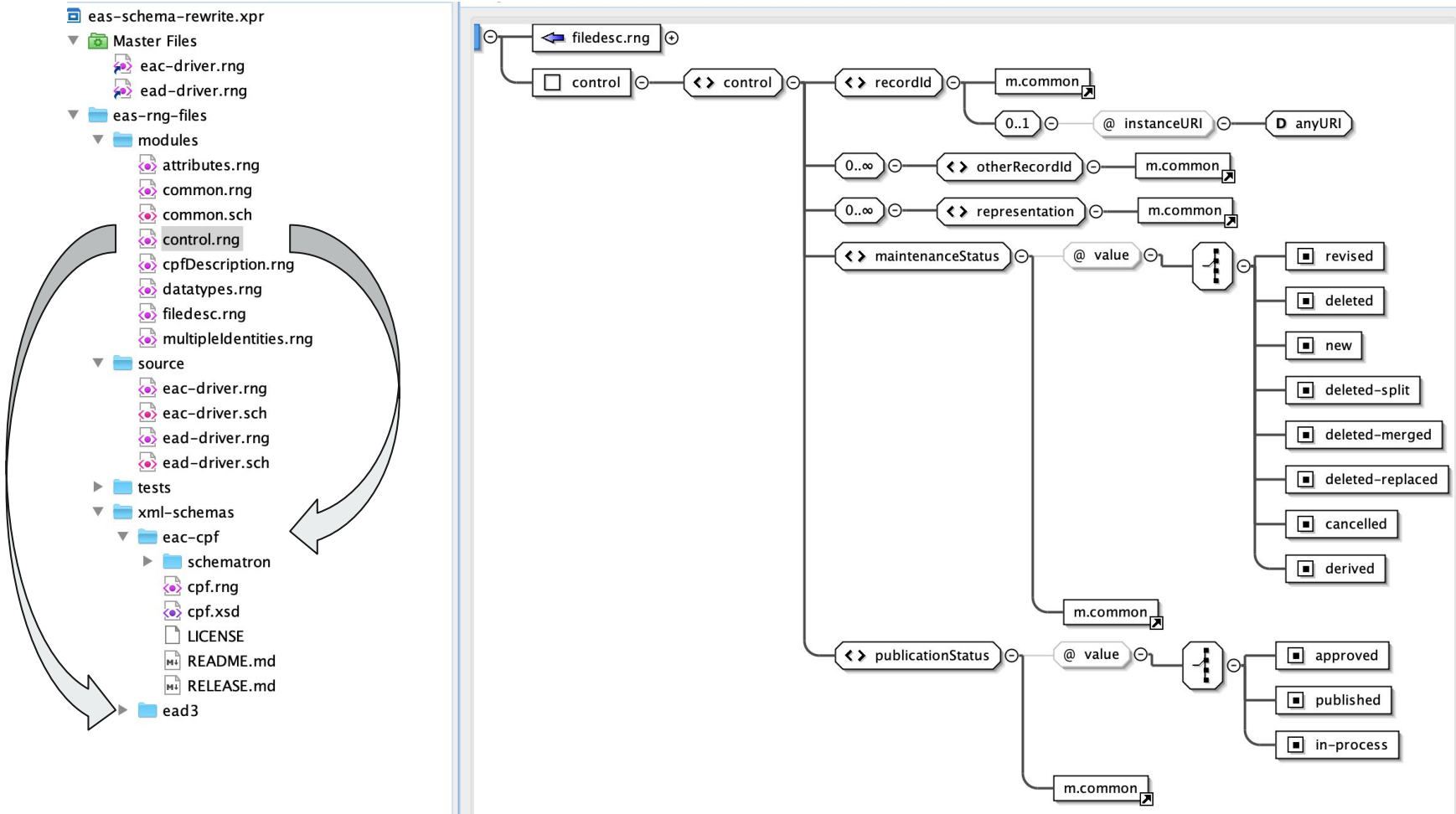
Mark

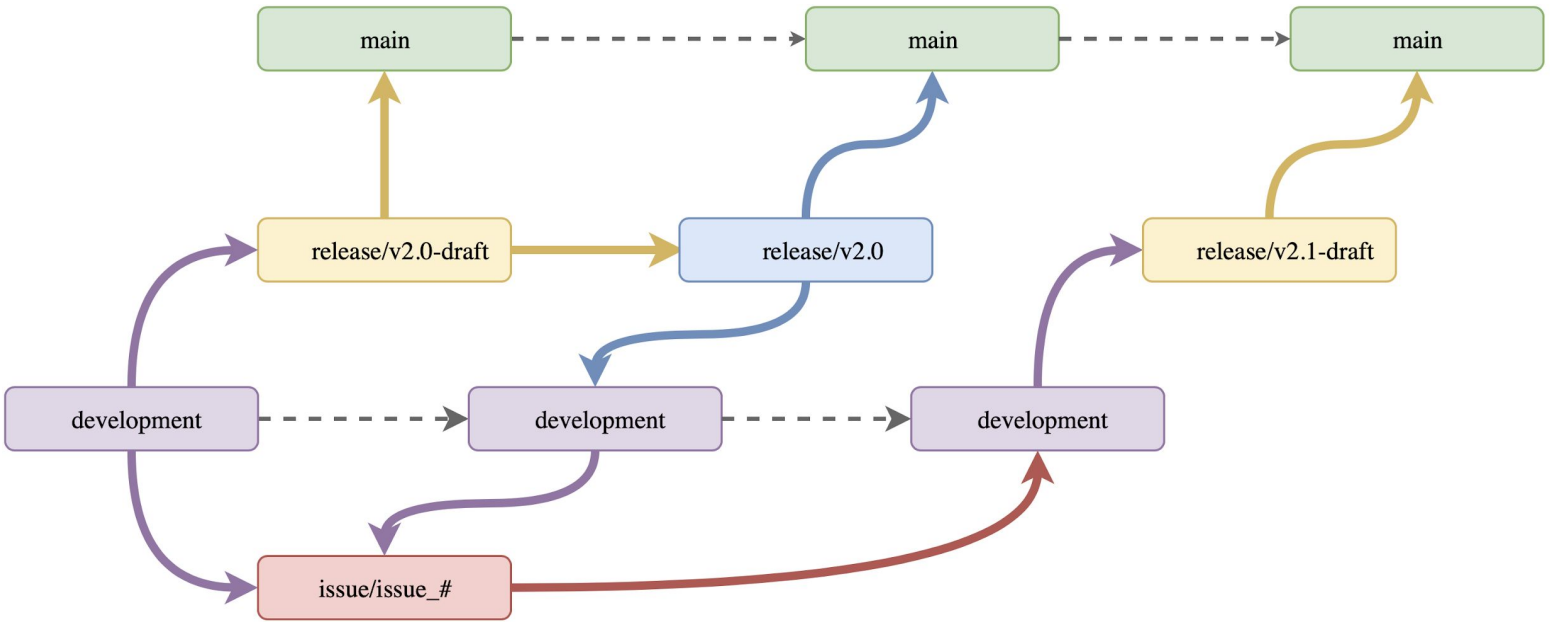
One schema, each element defined once

General direction

- One GitHub repository (eas-schemas) for developing EAD/EAC, etc.
  - The new repo will not be created until after the release of EAC-CPF 2.0, but we will follow the rest of the approach now.
- Shared attributes, elements and types will be defined once, in modularized Relax NG files
  - e.g. modules/control.rng
- Improved Git flow
  - Two constant branches: “main” and “development”
  - Releases (e.g. EAD3 1.1.1) will be created from “**release/***” branches, which break off of “development”
  - Set up GitHub actions
    - to test any pull requests to “main” and “development”
    - to automate the schema publication process so that the schemas will be generated in a containerized environment rather than locally
- Any differences between the schemas (**e.g. if eac:biogHist and ead3:biogHist differ**) will be annotated so that they can be handled during the publication process


eas-schema-rewrite.xpr
Master Files
eac-driver.rng
ead-driver.rng
eas-rng-files
modules
attributes.rng
common.rng
common.sch
control.rng
cpfDescription.rng
datatypes.rng
filedesc.rng
multipleIdentities.rng
source
eac-driver.rng
eac-driver.sch
ead-driver.rng
ead-driver.sch
tests
xml-schemas
eac-cpf
schematron
cpf.rng
cpf.xsd
LICENSE
README.md
RELEASE.md
ead3
filedesc.rng
control
control
recordId
m.common
0..1
@ instanceURI
D anyURI
0..∞
otherRecordId
m.common
0..∞
representation
m.common
maintenanceStatus
@ value
revised
deleted
new
deleted-split
deleted-merged
deleted-replaced
cancelled
derived
m.common
publicationStatus
@ value
approved
published
in-process
m.common


eas-schema-rewrite.xpr
Master Files
eac-driver.rng
ead-driver.rng
eas-rng-files
modules
attributes.rng
common.rng
common.sch
control.rng
cpfDescription.rng
datatypes.rng
filedesc.rng
multipleIdentities.rng
source
eac-driver.rng
eac-driver.sch
ead-driver.rng
ead-driver.sch
tests
xml-schemas
eac-cpf
schematron
cpf.rng
cpf.xsd
LICENSE
README.md
RELEASE.md
ead3
filedesc.rng
control
control
recordId
m.common
0..1
@ instanceURI
D anyURI
0..∞
otherRecordId
m.common
0..∞
representation
m.common
maintenanceStatus
@ value
revised
deleted
new
deleted-split
deleted-merged
deleted-replaced
cancelled
derived
m.common
publicationStatus
@ value
approved
published
in-process
m.common


main
main
main
release/v2.0-draft
release/v2.0
release/v2.1-draft
development
development
development
issue/issue_#

Namespaces

History of EAS namespaces

EAD: Released in 1998. No namespace. Only a DTD

EAD 2002: Also a DTD, but for the RNG and XSD outputs, the first EAD namespace:
urn:isbn:1-931666-**22-9**

EAC-CPF: Released in 2010. No more DTD, but another ISBN for the namespace:
urn:isbn:1-931666-**33-4**

EAD3: Released in 2015. Another DTD, but also *two* new HTTP namespaces for the RNG and XSD

- http://ead3.archivists.org/schema/
- http://ead3.archivists.org/schema/undeprecated/

Versioning

For more info, see: http://www.xfront.com/Versioning.pdf

For a sensible approach, there are many examples, but let's look at ALTO:https://www.loc.gov/standards/alto/

Note the shared practice with Semantic Versioning for the schemas.

Example ALTO namespace, for major version 4: http://www.loc.gov/standards/alto/ns-v4#

Summary

For EAC-CPF 2.0, we need a new namespace.

Minor revisions do not require namespace changes, but major revisions do.

- E.g. eac:functions (1.0) and eac:functions (2.0) will *not* be the same, so the namespace must be different.

SAA owns the ISBN numbers assigned to EAD2002 and EAC-CPF

SAA controls the URLs assigned for both EAD3 namespaces

EAS Namespace Proposal

Let's take the opportunity to switch to a URI for the EAC namespace

Continue to have SAA be the entity that manages the namespace

Add "ns" to the URL rather than "schema" to indicate that the URL is a namespace (common convention)

Change urn:isbn:1-931666-33-4 → http://www.archivists.org/eac/ns/v2

With the EAD redesign, consider implementing a homogeneous namespace for all of EAS: http://www.archivists.org/eas/ns/v1

Continue hosting the schemas elsewhere, and do *not* ask SAA to redirect the namespace links (as is done with EAD3)

Casing of element/attribute tags

Mark

Results of the vote

- **camelCase**
- kebab-case
- lowercase

Modelling events

Gerhard

Entity-Relations: current state

EAC-CPF	EAD
CPF to CPF	Resource to Resource
CPF to Function	Resource to Function
CPF to Resource	Resource to CPF
CPF to Time	Resource to Time
CPF to Place	Resource to Place
CPF to Subject (Mandate, Function, Occupation, Legal Status)	Resource to Subject (Subject, Function, Occupation, Legal Status, Works)
CPF to Local Description	Resource to Works (Bibliography)
	Resource to Instance (dao)

Entity-Relations: EAC-CPF relations

- New wrapper element //relation replaces: //cpfRelation, //resourceRelation, and //functionRelation
- New set of child elements to structure the description of relations:
  - //targetEntity
  - //relationType
  - //targetRole
  - //date | //dateRange | //dateSet
  - (//places)
- New optional attributes to identify vocabularies/ontologies:
  - @vocabularySource
  - @vocabularySourceURI
  - @valueURI

Entity-Relations: EAC-CPF relations

```
<relations>
    <relation id="1" sourceLink="#43" maintenanceLink="#43">
        <targetEntity targetType="person" vocabularySource="GND" vocabularySourceURI="http://d-nb.info/gnd/"
        valueURI="http://d-nb.info/gnd/1190671301">
            <part localType="family">Arendt</part>
            <part localType="given">Paul</part>
            <part localType="dates">1873-1913</part>
        </targetEntity>
        <relationType vocabularySource="GND" vocabularySourceURI="https://d-nb.info/standards/elementset/gnd#"
        valueURI="https://d-nb.info/standards/elementset/gnd#Family" localType="text"
        lang="en">Family</relationType>
        <relationType vocabularySource="GND" vocabularySourceURI="https://d-nb.info/standards/elementset/gnd#"
        valueURI="https://d-nb.info/standards/elementset/gnd#Family" localType="code">bezf</relationType>
        <targetRole vocabularySource="schema.org" vocabularySourceURI="http://schema.org"
        valueURI="http://schema.org/parent" localType="text" lang="en">parent</targetRole>
        <targetRole vocabularySource="schema.org" vocabularySourceURI="http://schema.org"
        valueURI="http://schema.org/parent" localType="code">par</targetRole>
        <date>...</date>
        <place>...</place>
        <descriptiveNote>...</descriptiveNote>
    </relation>
</relations>
```

Entity-Relations: EAC-CPF events

Berlin Meeting in March 2020

- Encoding of events only in EAD3 and only as child of //chronitem*
- Events identified with an EAC-CPF or EAD entity should be examined in more detail in context of:
  - Places and Time an event took place (marriage, conference, …)
  - Records in Context core entity “activity” (events, processes, functions, …)
  - Interoperability with authority files covering events like conferences

Entity-Relations: EAC-CPF events

Chicago Meeting in August 2020

- What should we do?
  - Task 1: Examining the nature of events in context of EAC-CPF/EAD
  - Task 2: Decision making on encoding approaches of events in EAS
  - More?

Guideline for community feedback

Karin

A council created guideline

See you tomorrow!

Tuesday 4th of August
2.00 pm PDT / 3.00 pm MDT / 4.00 pm CDT / 5.00 pm EDT /
10.00pm BST / 11.00 pm CEST/ 7.00am Wednesday AEST